

1657
1657
OIPE II 3
RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697DATE: 02/16/2000
TIME: 02:14:57

INPUT SET: S34756.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

MAY 01 2000
TECH CENTER 1600/2901

SEQUENCE LISTING

ENTERED

(1) General Information:

(i) APPLICANT: Choulika, Andre
Perrin, Arnaud
Dujon, Bernard
Nicolas, Jean-Francois

(ii) TITLE OF INVENTION: Nucleotide Sequence Encoding the Enzyme
I-SCEI and the Uses Thereof

(iii) NUMBER OF SEQUENCES: 52

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunner
(B) STREET: 1300 I Street, N.W.
(C) CITY: Washington
(D) STATE: D.C.
(E) COUNTRY: USA
(F) ZIP: 20005-3315

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: Unassigned
(B) FILING DATE: 25-JAN-2000
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/244,130
(B) FILING DATE: 04-FEB-1999

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/119,024
(B) FILING DATE: 20-JUL-1998

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/336,241
(B) FILING DATE: 07-NOV-1994

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697MAY 01 2000
TECH CENTER 1600/2900
DATE: 02/16/2000
TIME: 02:14:57

INPUT SET: S34756.raw

47 (vii) PRIOR APPLICATION DATA:
48 (A) APPLICATION NUMBER: US 07/971,160
49 (B) FILING DATE: 05-NOV-1992
50
51 (vii) PRIOR APPLICATION DATA:
52 (A) APPLICATION NUMBER: US 07/879,689
53 (B) FILING DATE: 05-MAY-1992
54
55 (viii) ATTORNEY/AGENT INFORMATION:
56 (A) NAME: Meyers, Kenneth J.
57 (B) REGISTRATION NUMBER: 25,146
58 (C) REFERENCE/DOCKET NUMBER: 3495-0111-11
59
60 (ix) TELECOMMUNICATION INFORMATION:
61 (A) TELEPHONE: 202-408-4000
62 (B) TELEFAX: 202-408-4400.
63
64
65 (2) INFORMATION FOR SEQ ID NO:1:
66
67 (i) SEQUENCE CHARACTERISTICS:
68 (A) LENGTH: 714 base pairs
69 (B) TYPE: nucleic acid
70 (C) STRANDEDNESS: single
71 (D) TOPOLOGY: linear
72
73 (ii) MOLECULE TYPE: DNA (genomic)
74
75
76
77 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
78
79 ATGCATATGA AAAACATCAA AAAAAACCAG GTAATGAACC TCGGTCCGAA CTCTAAACTG 60
80 CTGAAAGAAT ACAAATCCCA GCTGATCGAA CTGAACATCG AACAGTTCGA AGCAGGTATC 120
81 GGTCTGATCC TGGGTGATGC TTACATCCGT TCTCGTGATG AAGGTAAAAC CTACTGTATG 180
82 CAGTTCGAGT GGAAAAACAA AGCATACATG GACCACGTAT GTCTGCTGTA CGATCAGTGG 240
83 GTACTGTCCC CGCCGCACAA AAAAGAACGT GTTAACCACC TGGGTAACCT GGTAATCACC 300
84 TGGGGCGCCC AGACTTTCAA ACACCAAGCT TTCAACAAAC TGGCTAACCT GTTCATCGTT 360
85 AACAAACAAA AAACCATCCC GAACAACCTG GTTGAAACT ACCTGACCCC GATGTCTCTG 420
86 GCATACTGGT TCATGGATGA TGGTGGTAAA TGGGATTACA ACAAAAATC TACCAACAAA 480
87 TCGATCGTAC TGAACACCCA GTCTTTCACT TTCGAAGAAG TAGAATACCT GGTAAAGGGT 540
88 CTGCGTAACA AATTCCAAC GAACTGTTAC GTAAAAATCA ACAAAAACAA ACCGATCATC 600
89 TACATCGATT CTATGTCTTA CCTGATCTTC TACAACCTGA TCAAACCGTA CCTGATCCCC 660
90
91
92
93
94
95
96
97
98
99

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697DATE: 02/16/2000
TIME: 02:14:57

INPUT SET: S34756.raw

100 CAGATGATGT ACAAAGTGCC GAACACTATC TCCTCCGAAA CTTTCCTGAA ATAA 714
101
102
103 (2) INFORMATION FOR SEQ ID NO:2:
104
105 (i) SEQUENCE CHARACTERISTICS:
106 (A) LENGTH: 237 amino acids
107 (B) TYPE: amino acid
108 (D) TOPOLOGY: linear
109
110 (ii) MOLECULE TYPE: peptide
111
112
113 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
114
115 Met His Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro
116 1 5 10 15
117 Asn Ser Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn
118 20 25 30
119 Ile Glu Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr
120 35 40 45
121 Ile Arg Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp
122 50 55 60
123 Lys Asn Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp
124 65 70 75 80
125 Val Leu Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn
126 85 90 95
127 Leu Val Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn
128 100 105 110
129 Lys Leu Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn
130 115 120 125
131 Asn Leu Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe
132 130 135 140
133 Met Asp Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys
134 145 150 155 160
135 Ser Ile Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr
136 165 170 175
137 Leu Val Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys
138 180 185 190
139 Ile Asn Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu
140
141
142
143
144
145
146
147
148
149
150
151
152

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697DATE: 02/16/2000
TIME: 02:14:58INPUT SET: S34756.raw
205

153 195 200
154 Ile Phe Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr
155 210 215 220
156
157 Lys Leu Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys
158 225 230 235
159
160

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 722 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

172
173
174
175 AAAAATAAAA TCATATGAAA AATATTAAAA AAAATCAAGT AATCAATCTC GGTCCTATTT 60
176
177 CTAAATTATT AAAAGAATAT AAATCACAAT TAATTGAATT AAATATTGAA CAATTGGAAG 120
178
179 CAGGTATTGG TTTAATTTTA GGAGATGCTT ATATTCGTAG TCGTGATGAA GGTAAACTT 180
180
181 ATTGTATGCA ATTTGAGTGG AAAATAAGG CACATGGA TCATGTATGT TTATTATATG 240
182
183 ATCAATGGGT ATTATCACCT CCTCATAAAA AAGAAAGAGT TAATCATTTA GGTAATTTAG 300
184
185 TAATTACCTG GGGAGCTCAA ACTTTTAAAC ATCAAGCTTT TAATAAATTA GCTAACTTAT 360
186
187 TTATTGTAAA TAATAAAAAA CTTATTCCTA ATAATTTAGT TGAATATTAT TTAACACCTA 420
188
189 TGAGTCTGGC ATATTGGTTT ATGGATGATG GAGGTAAATG GGATTATAAT AAAAATTCTC 480
190
191 TTAATAAAAAG TATTGTATTA AATACACAAA GTTTACTTT TGAAGAAGTA GAATATTTAC 540
192
193 TTAAAGGTTT AAGAAATAAA TTTCAATTAA ATTGTTATGT TAAAATTAAT AAAAATAAAC 600
194
195 CAATTATTTA TATTGATTCT ATGAGTTATC TGATTTTTTA TAATTTAATT AACCTTATT 660
196
197 TAATTCCTCA AATGATGTAT AACTGCCTA ATACTATTTT ATCCGAAACT TTTTAAAT 720
198
199 AA 722

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 235 amino acids
(B) TYPE: amino acid

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/492,697DATE: 02/16/2000
TIME: 02:14:58

INPUT SET: S34756.raw

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro Asn Ser
1 5 10 15

Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn Ile Glu
20 25 30

Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr Ile Arg
35 40 45

Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp Lys Asn
50 55 60

Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp Val Leu
65 70 75 80

Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn Leu Val
85 90 95

Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn Lys Leu
100 105 110

Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Leu Ile Pro Asn Asn Leu
115 120 125

Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe Met Asp
130 135 140

Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Leu Asn Lys Ser Ile
145 150 155 160

Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Cys Tyr Leu Val
165 170 175

Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys Ile Asn
180 185 190

Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile Phe

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/09/492,697

DATE: 02/16/2000
TIME: 02:14:58

INPUT SET: S34756.raw

Line	Error	Original Text
31	Wrong application Serial Number	(A) APPLICATION NUMBER: Unassigned